T-1.43-RT

PROJECT THEMIS

Systems Research Center
Industrial & Systems Engineering Department
University of Florida
Gainesville

26

RELIABILITY MODELS FOR COMPONENTS
WITH CORRELATED FAILURES

Technical Report No. 26

SUBMITTED BY:

RL Patterson

R. L. Patterson
Principal Investigator

APPROVED BY:

RL Patterson

R. L. Patterson
Director, Project THEMIS

Department of Industrial and Systems Engineering
The University of Florida
Gainesville, Florida 32601

November 1969

This research was performed under Project THEMIS, ARO-D Contract No.
DAH C04 68C0002.

ABSTRACT

Standard reliability prediction formulas for multi-component systems make the assumption of statistical independence of the conditions of the components after a fixed period of time in a hypothesized environment. Although laboratory and field experience shows that this assumption is not always valid it persists as a basis for reliability modeling among practitioners of the art. The most relevant reason for this is that engineers generally are not noted for their knowledge of the mathematics of probability and one soon discovers that the assumption of statistical independence among components usually leads to the simplest mathematics. A second reason why the independence assumption is so predominant is that engineers and analysts are not clear on what alternatives should be pursued or even what the alternatives might be so there would be little point in making assumptions the modeling implications of which are simply not understood. There is another justification for the statistical independence assumption, that being the fact that such an assumption yields models that can provide bounds on system reliabilities. Analytical models which attempt to account for environmental effects on component failure rates show that the statistical independence assumption often leads to gross overestimates or underestimates of system reliability. Evans [1] argues that it is almost never the case that this assumption is correct, but also points out that other models which attempt to quantify or correct this error have their own limitations. If, on the other hand, it could be shown that an alternative model even though an approximation is indeed a closer representation to the true state of the world and is at the same time practical to use in a computational sense then it should be used. The question of validity lies with the basic assumptions underlying the model and not with the mathematics itself.

If a set of assumptions alternative to the assumption of statistical independence are accepted as more closely characterizing an environmental situation then the resulting mathematical predictions would be tentatively preferred, subject to experimental verification.

This paper reviews some previous explorations of the question of statistical dependence of operating components and the effects on system reliability and presents some results not previously worked out. In particular there is shown to exist a connection among three models that do not require the assumption of statistical independence of components. Under certain conditions these models can generate the same reliability prediction even though they appear to be derived under different assumptions.

TABLE OF CONTENTS

CHAPTER I

DEFINITIONS

For what follows it is necessary to distinguish among the notions of the physical dependence of two events (or two random variables), the conditional stochastic independence of two events given a third event, and the unconditional stochastic independence of two events.

<u>Definition 1</u> If two events A and B are unconditionally statistically independent then $P(AB) = P(A)\ P(B)$.

The question of the conditional statistical independence of two events A and B given a third event E arises when components are placed into a random operating environment. The events A and B represent the conditions of two components after a suitable period of time when operated in the presence of a randomly selected environmental condition E.

<u>Definition 2</u> If two events A and B jointly conditioned with respect to an event E are statistically independent under E then

$$P(AB/E) = P(A/E)\ P(B/E).$$

Let $\mathcal{E} = \{E_i : i \in I\}$ be a collection of mutually exclusive and exhaustive events called an <u>environmental profile</u>. If A and B are statistically independent under the environmental profile $\mathcal{E}$, then $P(A,B/E_i) = P(A/E_i)\ P(B/E_i)$ for all $E_i \in \mathcal{E}$. Thus if AB occurs only in the presence of an $E_i \in \mathcal{E}$ and A and B are statistically independent under all $E_i \in \mathcal{E}$, then

$$\int_{E_i} P(AB/E_i)dP(E_i) = \int_{E_i} P(A/E_i)\ P(B/E_i)dP(E_i).$$

If, in addition, A and B are unconditionally statistically independent then

$$\int_{E_i} P(A/E_i)\, P(B/E_i)\,dP(E_i) = \int_{E_i} P(A/E_i)\,dP(E_i) \int_{E_i} P(B/E_i)\,dP(E_i) = P(A)\, P(B).$$

The intuitive notion of the physical independence of two components is that the condition of each component neither affects nor is affected by the condition of the other. Following Evans (p. 348) this idea is restated quantitatively as follows.

<u>Definition 3</u> Two events A and B are physically independent under an environmental profile $\mathcal{E}$ if and only if they are statistically independent under $\mathcal{E}$, i.e., that

$$P(A,B/E_i) = P(A/E_i)\, P(B/E_i) \qquad \text{for all } E_i \in \mathcal{E}^{*}.$$

Thus, if the malfunction of component i necessarily destroys component j, intuition states that they are certainly not physically independent under such an environmental circumstance. If such an event were included in $\mathcal{E}$ as E_i with positive probability mass assignment then a calculation will also show that the events A and B representing the conditions of the components are not statistically independent under E_i and hence not physically independent under E_i.

The next definition characterizes the notion of "associated random variables" introduced by Esary, Proschan, and Walkup [2] in 1966. This definition is relevant to what follows because Pollyak [3] showed in 1962 that under a random environmental profile as specified by Definition 2 the individual component reliability functions are random variables which under certain conditions are associated in the sense of Esary.

<u>Definition 4</u> Random variables $X_1, X_2, \ldots, X_n$ are <u>associated</u> if $\text{Cov}\,[U(\overline{X}), V(\overline{X})] \geq 0$ for all pairs U,V of bounded continuous nondecreasing

*It is also the case that if $P(AB/E_i) = P(A/E_i)P(B/E_i)$ and $P(A/E_i) = P(A)$ for all E_i then $P(AB) = P(A)P(B)$.

functions where $\overline{X} = (X_1, X_2, \ldots, X_n)$. Esary, et al., showed that association has the following properties:

Property 1 Any subset of a set of associated random variables is also a set of associated random variables.

Property 2 If two sets of associated random variables are independent of one another then their union is a set of associated random variables.

Property 3 The set consisting of a single random variable is associated.

Property 4 If $X_1, X_2, \ldots, X_N$ are associated then any set of nondecreasing functions $S_1(\overline{X}), S_2(\overline{X}), \ldots, S_M(\overline{X})$ are associated.

The concept of equally ordered functions will be of use and is specified by the following definition.

Definition 5 Let $F(t,\overline{X})$ and $G(t,\overline{X})$ be functions of a parameter t and a vector argument $\overline{X} = (X_1, \ldots, X_k)$ where X_i is a real number. F and G are said to be equally ordered with respect to the variables $X_1, \ldots, X_k$ if for any values of $\overline{X}_1$ and $\overline{X}_2$ of the vector argument $\overline{X}$ the inequality

$$[F(t,\overline{X}_1) - F(t,\overline{X}_2)][G(t,\overline{X}_1) - G(t,\overline{X}_2)] \geq 0$$

is satisfied. Monotonic nonincreasing or nondecreasing functions of a single scalar argument defined on the real line belong to this class as well as certain convex functions.

If F and G in Definition 5 are functions of a random vector $\overline{X} = (X_1, \ldots, X_k)$, then the variables $X_1, \ldots, X_k$ are associated in the sense of Definition 4.

CHAPTER II

THE ENVIRONMENTAL PROFILE MODEL OF RELIABILITY

It is evident that most components of any system function within an environment that is "uncertain" in the sense that the intensity of its defining variables exhibit statistical fluctuation over time. Examples of environmental variables are temperature, pressure, humidity, vibration frequency, salt content, and radioactive emission level.

Let the environment be characterized by a random vector $\overline{X} = (X_1, \ldots , X_k)$ where X_i represents the i-th environmental variable defined over a suitable range of values. The environmental profile is specified by the set

$$\mathcal{E} = \{(X_1, \ldots , X_k): \ \alpha_i < X_i < \beta_i, \ \ i = 1, \ldots , k\} .$$

Let the joint probability density of $\overline{X}$ be $\omega(\overline{X}) = \omega(X_1, \ldots , X_k)$

Let $R_i(t,\overline{X})$ be the conditioned reliability of the i-th system component where t denotes time and $\overline{X}$ indicates that reliability is a function of (or conditioned upon) $\overline{X}$. Pollyak's model of component reliability is obtained as an average over the environmental profile $\mathcal{E}$:

$$R(t) = \int_{\overline{X}} R(t,\overline{X}) \, d\omega(\overline{X}) \quad , \quad \overline{X} \in \mathcal{E} . \tag{1}$$

The reliability of n components in series, assuming mutual statistical independence of the components under $\mathcal{E}$ is

$$R_N(t) = \int_{\overline{X}} \prod_1^N R_i(t,\overline{X}) \, d\omega(\overline{X}) \quad , \quad \overline{X} \in \mathcal{E} . \tag{2}$$

$R_i(t,\overline{X})$ is usually assumed to be monotonic nonincreasing in t but what is more interesting is the effect on $R_N(t)$ of averaging the product of the component reliabilities over the sample space of environmental possibilities. Pollyak

answered this question by proving the following theorem.

Theorem 1 Let $R_i(t,\overline{X})$ and $R_j(t,\overline{X})$ be functions of the k-dimensional vector argument $\overline{X}$ that are equally ordered in the sense of Definition 5. Let $\overline{X}$ belong to a set $\mathcal{E}$ and let $\omega(\overline{X})$ be a weight function defined on $\mathcal{E}$ which is everywhere nonnegative and $\int_{\overline{X}} \omega(\overline{X}) d\overline{X} = 1$.

Then

$$\int_{X\varepsilon\mathcal{E}} R_i\, R_j\, \omega(\overline{X}) d\overline{X} - \int_{X\varepsilon\mathcal{E}} R_i\, \omega(\overline{X}) d\overline{X} \int_{X\varepsilon\mathcal{E}} R_j\, \omega(\overline{X}) d\overline{X} \geq 0 \tag{3}$$

Equation (3) implies that $Cov(R_i,R_j) \geq 0$. Consequently the effect of imbedding component reliabilities within a random environment and averaging over the environmental possibilities with respect to a distribution to obtain the unconditional reliability is the following:

(a) For serial systems reliability is at least as high as the value predicted by making the assumption of unconditional statistical independence with respect to the failures of the components.

(b) For parallel systems reliability is at least as low as the value predicted by making the assumption of unconditional statistical independence among the conditions of the components.

In other words, if the environmental profile model as defined by Equation (2) is a more realistic reflection of physical reality and if component failure laws satisfy conditions of Theorem 1 then standard prediction formulas for serial reliability give underestimates and for parallel system reliability they give optimistic results. (This interpretation follows since pairs of products of equally ordered functions are equally ordered. For an alternate justification see Theorem 2.)

It seems plausible that the assumption of monotonicity of the component failure law with respect to the degree of severity of the environment $\overline{X}$

would be satisfied in most problem situations. The environmental variables would be represented either directly or implicitly in terms of the moments of the $R_i(t,\overline{X})$. For example if $R_i(t,\overline{X}) = e^{-\theta_i t}$ then θ_i depends upon $\overline{X}$ perhaps as a linear regression upon $(X_1, \ldots, X_k)$. Equation (3) does not imply that one needs to imbed components in a random operating environment which produces a positive correlation among component conditions in order to increase system reliability. To attempt to do so to an extent greater than what is usually the case anyway may impair system reliability. Equation (3) does imply however that components operating in a random environment and having failure laws that are monotonically related to the environmental parameters tend to fail together or function properly together. That is, if the environmental model applies, knowledge that a component has either failed (or survived) increases the betting odds that another component also failed (or survived).

The fact that the environmental profile model is a more accurate reflection of reality means not that correlated conditions among components produces higher reliabilities, but rather that the assumption of unconditional statistical independence of failures was not realistic to begin with.

It is possible, on the other hand, to conceive of a situation in which the environmental profile is altered by placing components in a state of physical dependence. The effects are reflected in correlated failures and serial system reliability is actually increased over what it would be had the components not been placed in such a state of mutual phys cal dependence. An example is the case where two transistors are mounted on a common heat sink. The assumption of statistical independence of components under the environmental profile would not be justified and Equation (2) would have to be modified. Thus, correlations among component conditions can be present

with or without the assumption of their statistical independence under an environmental profile but the interpretation of these correlations depends upon which conditions hold.

A second order approximation of the difference between the reliability estimates given by Equation (2) and the model $R(t) = (p(t))^N$ can be obtained by the use of Taylor's series. When the environment is specified in terms of a scalar x the difference is calculated as follows for an N component serial system.

Let $R_N(t,x) = (R(t,x))^N$ and let $\omega(x)$ be the probability density of the environmental variable x. Equation (2) gives the system reliability as

$$R_N(t) = \underset{x}{E}\,(R_N(t,x)) = \underset{x}{E}\,[(R(t,x))^N].$$

The Taylors series approximation is

$$R_N(t) \doteq (R(t,x)\,)^N\Big|_{x=\mu} + \frac{\sigma_x{}^2}{2}\;\frac{d^2}{dx^2}\;(R(t,x))^N\Big|_{x=\mu} =$$

$$= R^N(t,\mu) + \frac{\sigma_x{}^2}{2}\cdot N\cdot R^{N-2}(t,\mu)\cdot[(N-1)(R'(t,\mu))^2 + R(t,\mu)R''(t,\mu)]$$

where primes denote differentiation with respect to x.

The difference Δ in the estimates is

$$\Delta \doteq \frac{\sigma_x{}^2}{2}\cdot N\cdot R^{N-2}(t,\mu)\cdot[(N-1)(R'(t,\mu))^2 + R(t,\mu)R''(t,\mu)] \qquad (4)$$

PROOF OF AN IMPORTANT INEQUALITY

The conclusions drawn in the previous section depend in part upon the validity of Equation (3). This inequality for the general case follows from:

<u>Theorem 2</u> Let $\eta = (\eta_1, \eta_2, \ldots, \eta_k)$ be a vector of random variables having

finite moments and having a joint density function $\omega(\eta)$. η is assumed to define an environmental profile ξ. Define N characteristic random variables as follows:

$$X_i = \begin{cases} 1 & \text{if i-th component survives in } \xi \text{ for time t.} \\ 0 & \text{otherwise} \end{cases}$$

Let the X_i's be mutually independent under ξ.

Let $R_N(t) = P_r$ (N components survive ξ for time t).

Then

$$R_{h+k}(t) \geq R_h(t)\ R_k(t) \qquad \text{for } h, k \text{ positive integers.}$$

<u>Proof</u> By definition,

$$R_N(t) = E(X_1\ X_2\ \dots\ X_N) = \underset{\eta}{E}E(X_1\ \dots\ X_N/\eta) = \underset{\eta}{E}\{E(X_1/\eta)E(X_2/\eta)\dots E(X/\eta)\}$$

$$= \underset{\eta}{E}(R^N(t,\eta)) = \underset{\alpha}{E}\ \alpha^N \qquad \text{Where } \alpha = R(t,\eta) = E(X_i/\eta)$$

$$(0 \leq \alpha \leq 1).$$

Thus R_N is the N-th raw moment of some distribution defined on [0,1].

For any positive integers h and k and independent random variables α_1 and α_2 defined on [0,1] and having the distribution of α, it follows that

$$(\alpha_1{}^h - \alpha_2{}^h)(\alpha_1{}^k - \alpha_2{}^k) \geq 0$$

so that

$$E(\alpha_1{}^{h+k}) - 2E(\alpha_1{}^h\ \alpha_2{}^k) + E(\alpha_2{}^{h+k}) =$$

$$2E(\alpha^{h+k}) - 2E(\alpha^h)E(\alpha^k) \geq 0\ .$$

Thus $E(\alpha^{h+k}) \geq E(\alpha^h)E(\alpha^k)$.

Rewriting the last inequality in terms of reliabilities

$$R_{h+k}(t) \geq R_h(t)\ R_k(t)$$

Esary et al. proved a theorem analogous to Theorem 2 in terms of associated random variables $\eta_1, \ldots, \eta_k$ as well as proving another inequality of greater generality.

CHAPTER III

A STRESS-STRENGTH MODEL OF RELIABILITY

The environmental variables model has appeared in the literature in a somewhat different form. Lloyd and Lipow [4] as well as others have considered the "chain" or "weakest link" model, also called the "stress-strength" model. A generalized version of this model for the reliability of an n-component serial system follows.

Components are assumed to function in a random environment characterized by a sequence of stress "pulses", random in number over a time interval of fixed length and each following some distribution of stress intensity. Thus the environment is defined in terms of a discrete variable,

$N(t)$ = number of stress pulses in a time interval of length t

and a distribution of stress intensity

$F(x) = P_r$ (a randomly selected stress pulse does not exceed x in intensity)

Each component is assumed to possess a "strength" selected at random from a distribution $G(x)$. Reliable operation of the n-component system over a time interval of length t occurs whenever the minimum component strength exceeds the maximum stress intensity occurring during the period.

Let Y_n denote the maximum stress intensity in a time interval of length t given that n stress pulses occur.

Then, assuming independence of the pulse intensities,

$$P(Y_n \leq x) = (F(x))^n \qquad (n = 0, 1, 2, \ldots)$$

Assuming $N(t)$ and Y_n to be statistically independent,

$$P(N(t) = n, Y_n \leq x) = P(N(t) = n)(F(x))^n. \qquad (n = 0, 1, \ldots)$$

The marginal distribution of the maximum stress intensity in a time interval of length t, $Y(t)$, is therefore

$$P(Y(t) \leq x) = \sum_{n=0}^{\infty} P(N(t) = n)(F(x))^n .$$

For example, assume that $\{N(t); t \geq 0\}$ is a stationary Poisson process of intensity λ.

Then

$$P(Y(t) \leq x) = \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-\lambda t}(F(x))^n$$

$$= e^{-\lambda t(1-F(x))} \sum_{n=0}^{\infty} \frac{(\lambda t\ F(x))^n}{n!} e^{-\ t\lambda F(x)} .$$

Thus,

$$H(x;t) \doteq P(Y(t) \leq x) = e^{-\lambda t(1-F(x))} \tag{6}$$

If each component strength follows the distribution $G(x)$ and component strengths are independent, the distribution function $G_n(x)$ of the minimum component strength $X_{(n)}$ is therefore

$$G_n(x) = 1 - P(X_{(n)} > x) = 1 - (1 - G(x))^n.$$

The n-component serial reliability is defined as

$$R_n(t) = P(X_{(n)} > Y(t)) = \int_0^{\infty} (1 - G_n(x))dH(x;t) \tag{7}$$

Thus, in the special case given by Equation (6) and by making the plausible assumption that $F(0) = 0$,

$$R_n(t) = e^{-\lambda t} + \int_{0+}^{\infty} (1 - G(x))^n d\ e^{-\lambda t(1-F(x))} \tag{7'}$$

The similarity between Equation (7) and Equation (2) is obvious. Recalling that R_n from Equation (2) can be interpreted as the n-th raw moment of a distribution defined on [0,1] it is also clear that $R_n(t)$ from Equation (7) can be interpreted in the same manner. The model defined by Equation (7) can be generalized to the case of k distinct stress and strength variables.

CHAPTER IV

A MULTIVARIATE BERNOULLI MODEL OF RELIABILITY

A somewhat different approach toward accounting for dependencies among component conditions was taken by Patterson and Khanna [5] in which the basic result is stated as

Theorem 3 Let $X_1, \ldots, X_n$ be n Bernoulli random variables each having parameter p and possessing a joint probability function $p(X_1, X_2, \ldots, X_3)$. Let all pairs X_i, X_j have linear correlation ρ. Let all conditioned pairs $(X_{i_1}, X_{i_2} | X_{i_3} = 1, \ldots, X_{i_n} = 1)$ also have linear correlation ρ.

Then

$$P(X_1 = 1, \ldots, X_n = 1) = \prod_{i=1}^{n} [1 - (1-\rho)^{i-1}(1-p)] \qquad (0 \le \rho \le 1) \quad (n = 1,2,3,\ldots) \tag{8}$$

If $P(X_i = 1)$ denotes the reliability of the i-th component then the model given by Equation (8) gives the n-component serial reliability of a system in which component conditions exhibit a statistical dependence. The reliability of an n component parallel system can be obtained from (8) by the expression

$$1 - P(X_1 = 0, X_2 = 0, \ldots, X_n = 0) = \prod_{i=1}^{n} [1 - (1-\rho)^{i-1}p] \tag{8'}$$

From Equation (8) it is seen that

$$R_1^n \le R_n \le R_1 \qquad (0 \le \rho \le 1)$$

where $R_n = P(X_1 = 1, \ldots, X_n = 1)$.

Geometrically the probabilities $P(X_1, \ldots, X_n)$ are mass assignments to the vertices of an n-dimensional unit cube that sum to unity. For the $n = 2$ case Equations (8) and (2) are equivalent if

$$P(X_i = 1) = \int_{\overline{X}} R(t,\overline{X})\omega(\overline{X})d\overline{X} \qquad (i = 1, 2).$$

Nothing is assumed, however, in the derivation of Equation (8) about the cause mechanism either with respect to the environment or possible physical connections among the components that have the effect of producing correlations between component failures. The assumption is, simply, that it is possible to express total serial system reliability as a product of conditional probabilities such that the probability of survival of a given component increases when it is known that increasing numbers of the other $n - 1$ components of the system survived. Since each conditional probability is of the form

$$P(X_{i_{k+1}} = 1 | X_{i_1} = 1, X_{i_2} = 1, \ldots, X_{i_k} = 1) = 1 - (1-\rho)^k q$$

the additional knowledge that the (k+1) st component survived increases the conditional probability that one of the remaining $n-(k+1)$ components survived by the amount $\rho q(1-\rho)^k$.

Although the multivariate Bernoulli model given by Equation (8) represents a different approach to the problem the sequence

$$\{R_n\} = \{\prod_{i=1}^{n} (1-\beta^{i-1}q)\} \quad (\beta = 1-\rho) \quad , \quad 0 \leq \beta \leq 1)$$
$$(p = 1-q)(n = 1,2, \ldots)$$

represents the moments of a probability distribution defined over the closed unit interval. This fact was demonstrated in a corollary proved by John Saw which depends upon a theorem due to F. Hausdorff and discussed by Feller [6].

The statements of the theorem and corollary follow.

Definition Let F be a probability distribution concentrated on the interval [0,1]. The k-th moment μ_k of F is

$$\mu_k = E(X^k) = \int_0^1 X^k \, dF.$$

Let $E(X^k(1 - X)) = -\Delta\mu_k$.

Then by induction

$$(-\Delta)^r\mu_k = E(X^k(1-X)^r).$$

Theorem

A sequence of numbers $\mu_0, \mu_1, \ldots$ represents the moments μ_k of some of some probability distribution F concentrated on [0,1] if and only if $(-\Delta)^r\mu_k \geq 0$, $\mu_0 = 1$.

Corollary

Let $0 < u_1 \leq u_2 < \ldots \leq 1$ be a nondecreasing set of positive values bounded above by unity and satisfying

$$(-\Delta)^r u_k \leq 0$$

and define $\mu_0 = 1$, $\mu_k = \prod_1^k u_i$.

Then $\mu_0, \mu_1, \ldots$ are the moments of some probability distribution concentrated on [0,1].

By letting $\mu_0 = 1$, $\mu_k = \prod_1^k (1 - \beta^{i-1}q)$

so that

$$u_i = 1 - \beta^{i-1}q \quad , \quad (i = 1, 2, \ldots)$$

then $\quad 0 < u_1 \leq u_2 \leq \ldots \leq 1$

and $\quad (-\Delta)^r u_k = -\beta^{k-1}(1 - \beta)^r q < 0.$

The conditions of the corollary are satisfied and therefore the sequence

$$\mu_0 = 1$$

$$\mu_k = \prod_{1}^{k} (1 - \beta^{i-1} q)$$

are the moments of a distribution F defined on [0,1]. Following Feller (p. 222, 223) the distribution F is constructed as follows.

Let

$$p_k^{(n)} = \binom{n}{k} (-\Delta)^{n-k} R_k$$

where

$$R_k = \prod_{1}^{k} (1 - \beta^{i-1} q).$$

Then

$$\begin{cases} p_{n+1}^{(n+1)} = R_{n+1} \\ p_k^{(n+1)} = q \dfrac{n+1}{n+1-k} \displaystyle\sum_{j=0}^{n-k} \binom{n}{j} \beta^{n-j} (1-\beta)^j p_k^{(n-j)} \qquad (k = 0, 1, \ldots, n) \end{cases}$$

Having computed the $p_k^{(n)}$ through this recursion formula for each x, define

$$F_n(x) = \sum_{k \le nx} p_k^{(n)}$$

and let

$$\lim_{n \to \infty} F_n(x) = F(x).$$

Then

$$\int_0^1 x^n dF(x) = R_n = \prod_{i=1}^{n} (1 - q\beta^{i-1})$$

Examples

1. $(\beta = 1 - \rho = 0)$

For $\beta = 0$, $R_0 = 1$, $R_k = p \qquad (k \ge 1)$

and

$$p_0^{(n)} = q, \quad p_k^{(n)} = 0 \qquad (0 < k < n)$$

$$p_n^{(n)} = p.$$

Then

$$F_n(x) = \sum_{k \leq nx} p_k^{(n)} = \begin{cases} q ; 0 \leq x < 1 \\ 1 ; x = 1 \end{cases}$$

and

$$F(x) = \begin{cases} q ; 0 \leq x < 1 \\ 1 ; x = 1 \end{cases}$$

2. $(\beta = 1)$

$$R_k = p^k \qquad k \geq 0$$

so that

$$p_k^{(n)} = \binom{n}{k} p^k (1 - p)^{n-k}$$

and

$$F_n(x) = \sum_{k \leq nx} \binom{n}{k} p^k (1 - p)^{n-k}$$

$$= 1 - \int_0^p \frac{u^{[nx]-1} (1 - u)^{n-[nx]}}{B([nx]; n - [nx] + 1)} du$$

where $[nx]$ is the largest integer satisfying $[nx] < nx$ and $B(r,s)$ is the Beta function with arguments r and s.

Thus

$$F(x) = \begin{cases} 0 ; x < p \\ 1 ; x \geq p \end{cases}$$

The preceding theory shows that the R_n given by Equations (2), (7), and (8) can each be interpreted as the n-th raw moments of a probability distribution defined on [0,1].

CONCLUSIONS

The models given by Equations (2), (7), and (3) represent three methods of eliminating the necessity of the assumption of unconditional statistical independence among components operating jointly in a common environment. It has been shown that R_n can be interpreted as the n-th raw moment of a probability distribution F(x) distributed over [0,1]. It was shown that the assumption of the unconditional statistical independence of component conditions results in an underestimate of serial system reliability and an overestimate of the reliability of a system of components in parallel.

The environmental profile vector $\overline{X} = (X_1, \ldots, X_k)$ constitutes a set of associated random variables whenever the reliabilities $p_i(\overline{X})$ and $p_j(\overline{X})$ change monotonically with respect to $\overline{X}$.

Finally, the environmental profile and Bernoulli models represent what might properly be called static models in the sense that none of the random variables involved represent stochastic processes. The stress-strength model is dynamic so far as the representation of stress is concerned.

The author expresses his gratitude to John Saw for helpful criticisms and suggestions in the preparation of this paper.

BIBLIOGRAPHY

[illegible] van[illegible] R. A., "Problems in Probability," Proceedings of the 1966 [illegible] Symposium on Reliability, pp. 347-53.

2. Esary, J. D., Proschan, F., and Walkup, D. W., "A Multivariate Notion of Association for General Random Variables," Mathematical Note No 490, Mathematics Research Laboratory, Boeing Scientific Research Laboratories, October 1966.

3. Pollyak, Y. G., "Errors in Reliability Prediction Caused by the Statistical Relation Between Component Failures," IEE Transactions.

4. Lloyd, D. K. and Lipow, M., "Reliability: Management, Methods, and Mathematics," Prentice-Hall, 1962, pp. 231-33.

5. [illegible]at[illegible]rson, R. L. and Khanna, K. M., "Reliability Models for Dependent Elements Using a Method of Correlations," 1967 (unpublished).

6. Feller, W., "An Introduction to Probability Theory and Its Applications," Vol. II, Wiley and Sons, p. 223.

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

University of Florida

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

Unclassified

3. REPORT TITLE

"Reliability Models for Components with Correlated Failures"

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (Last name, first name, initial)

Patterson, Richard L.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 1969	19	6

8a. CONTRACT OR GRANT NO.

DAH C04 68C 0002

b. PROJECT NO.

2T0 14501B81B

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Technical Report No. 26

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

ARO-Durham

13. ABSTRACT

Standard reliability prediction formulas for multi-component systems make the assumption of statistical independence of the conditions of the components after a fixed period of time in a hypothesized environment. Although laboratory and field experience shows that this assumption is not always valid it persists as a basis for reliability modeling among practitioners of the art. The most relevant reason for this is that engineers generally are not noted for their knowledge of the mathematics of probability and one soon discovers that the assumption of statistical independence among components usually leads to the simplest mathematics. A second reason why the independence assumption is so predominant is that engineers and analysts are not clear on what alternatives should be pursued or even what the alternatives might be so there would be little point in making assumptions the modeling implications of which are simply not understood. There is another justificantion for the statistical independence assumption, that being the fact that such an assumption yields models that can provide bounds on system reliabilities. Analytical models which attempt to account for environmental effects on component failure rates show that the statistical independence assumption often leads to gross overestimates or underestimates of system reliability. Evans [1] argues that it is almost never the case that this assumption is correct, but also points out that other models which attempt to quantify or correct this error have their own limitations. If, on the other hand, it could be shown that an alternative model even though an approximation is indeed a closer representation to the true state of the

DD FORM 1 JAN 64 1473

Security Classification

Abstract (continued)

world and is at the same time practical to use in a computational sense then it should be used. The question of validity lies with the basic assumptions underlying the model and not with the mathematics itself. If a set of assumptions alternative to the assumption of statistical independence are accepted as more closely characterizing an environmental situation then the resulting mathematical predictions would be tentatively preferred, subject to experimental verification.

This paper reviews some previous explorations of the question of statistical dependence of operating components and the effects on system reliability and presents some results not previously worked out. In particular there is shown to exist a connection among three models that do not require the assumption of statistical independence of components. Under certain conditions these models can generate the same reliability prediction even though they appear to be derived under different assumptions.